Abdominal Trauma Detection

Saurav Mistry Agun
School of Computer
Science and Engineering
Independent University Bangladesh
Dhaka, Bangladesh
firethedev@gmail.com

Arif Ul Hoque
School of Computer
Science and Engineering
Independent University Bangladesh
Dhaka, Bangladesh
arif.hoque123@gmail.com

Ricardo Dhrubo Hazra
School of Computer
Science and Engineering
Independent University Bangladesh
Dhaka, Bangladesh
ricardohazra@gmail.com

***Abstract*—Our paper presents a Convolutional Neural Network (CNN) based on the ResNet50 architecture, which is a deep residual network commonly used for image classification tasks and transfer learning. The model intends to detect Abdominal trauma which is a critical aspect of emergency medicine, as abdominal injuries can lead to life-threatening complications if not promptly diagnosed and treated. This study aims to develop a reliable method for detecting abdominal trauma using advanced imaging techniques and machine learning algorithms. By analyzing CT scans and other imaging modalities, our approach seeks to identify signs of organ damage, internal bleeding, and other abnormalities indicative of abdominal trauma. The proposed method has the potential to improve the accuracy and efficiency of abdominal trauma diagnosis, leading to better patient outcomes and reduced morbidity and mortality rates.**

Introduction

Abdominal trauma is a significant medical emergency often encountered in clinical practice, resulting from various causes such as motor vehicle accidents, falls, assaults, or industrial mishaps. Prompt and accurate diagnosis of abdominal trauma is crucial for appropriate patient management and improving clinical outcomes. In recent years, advancements in medical imaging technology, coupled with machine learning techniques, have shown promise in enhancing the detection and characterization of abdominal injuries. This report aims to investigate the application of advanced imaging modalities and machine learning algorithms for the detection and diagnosis of abdominal trauma. By leveraging state-of-the-art techniques, this study seeks to contribute to the development of more efficient and accurate methods for abdominal trauma assessment, ultimately improving patient care and prognosis.

Background

Abdominal trauma accounts for a significant portion of emergency department visits and poses a considerable risk to patient health and survival. The abdomen houses vital organs such as the liver, spleen, kidneys, and bowel, all of which are susceptible to injury from various traumatic events. Timely and accurate diagnosis of abdominal trauma is critical for guiding treatment decisions and minimizing morbidity and mortality rates. Traditional diagnostic approaches rely on physical examination, laboratory tests, and imaging studies such as ultrasound, computed tomography (CT), and magnetic resonance imaging (MRI). However, the interpretation of imaging findings can be challenging, requiring expertise and often leading to delays in diagnosis and treatment. In recent years, advancements in medical imaging technology and machine learning algorithms have shown promise in improving the accuracy and efficiency of abdominal trauma detection. By leveraging these advancements, this study aims to explore innovative approaches for enhancing the detection and diagnosis of abdominal trauma, with the ultimate goal of improving patient outcomes and healthcare delivery.

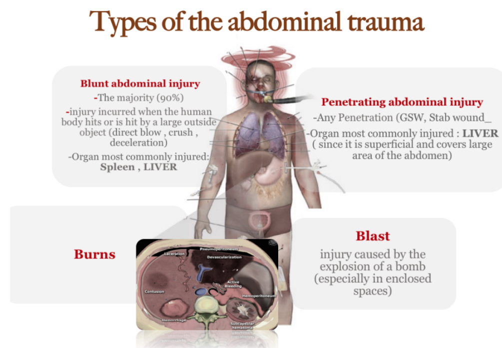

Fig. 1. Types of Abdominal Trauma

Methodology

This study employs a retrospective analysis of abdominal trauma cases, utilizing medical imaging data from computed tomography (CT) scans. A convolutional neural network (CNN) architecture, specifically ResNet50, is adopted for image analysis. Transfer learning techniques are utilized to fine-tune the pre-trained ResNet50 model for abdominal trauma detection. The dataset is divided into training and validation sets for model training and evaluation.

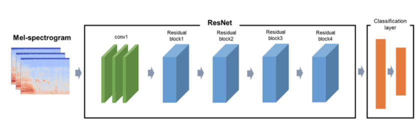

Fig. 2. Model Architecture

Model performance is assessed based on metrics such as accuracy, sensitivity, and specificity. Additionally, interpretability techniques are applied to understand the model's decision-making process. Overall, this methodology aims to develop an efficient and accurate approach for detecting abdominal trauma using advanced imaging and machine learning techniques.

DATASET

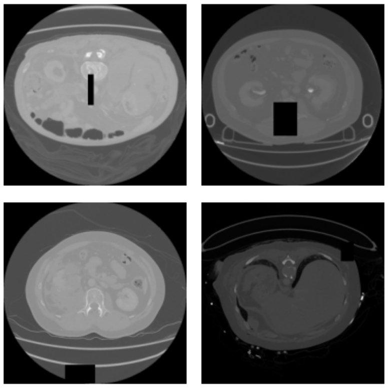

Fig. 3. Abdominal CT Scans

The figure shows some anonymous abdominal CT scans obtained from patients with suspected abdominal trauma. In addition to imaging data, the dataset includes patient-specific information such as clinical findings and outcomes. it indicates the overall correctness of the model's predictions across all classes, including bowel injury, extravasation injury, kidney conditions, liver conditions, and spleen conditions. These parameters provide valuable insights into the condition of different abdominal organs and contribute to the comprehensive analysis of abdominal trauma cases.

Data pre-processing techniques are applied to standardize image resolution and format for model compatibility. The dataset is divided into training and validation subsets to facilitate model training and evaluation. Ethical considerations are adhered to ensure patient privacy and confidentiality. Overall, the dataset serves as a valuable resource for developing and validating machine learning models for abdominal trauma detection.

A. *Problem Formulation*

The problem addressed in this project is the detection and diagnosis of abdominal trauma using medical imaging and machine learning techniques. Abdominal trauma, resulting from various causes such as accidents or injuries, poses a significant risk to patient health and survival. Prompt and accurate diagnosis of abdominal trauma is crucial for guiding treatment decisions and improving clinical outcomes.

Traditionally, the diagnosis of abdominal trauma relies on physical examination, laboratory tests, and imaging studies such as computed tomography (CT) scans. While imaging modalities provide valuable information, the interpretation of imaging findings can be challenging and time-consuming, requiring expertise and often leading to delays in diagnosis and treatment.

The goal of this project is to develop an efficient and accurate method for detecting and diagnosing abdominal trauma using advanced medical imaging techniques and machine learning algorithms. Specifically, the project aims to:

1) Utilize medical imaging data, particularly CT scans, to identify signs of abdominal injury, organ damage, and internal bleeding. Let X represent the set of input CT images, and Y represent the corresponding labels indicating the presence or absence of abdominal trauma.
2) Implement machine learning algorithms, including convolutional neural networks (CNNs), to analyze imaging data and extract relevant features indicative of abdominal trauma. Let $f(X;\theta)$ represent the learned mapping function parameterized by θ, which maps input images X to predicted labels $\hat{Y}$.
3) Train and validate the machine learning model using annotated datasets of abdominal trauma cases, incorporating patient-specific information and clinical findings. Let $D = \{(X_i, Y_i)\}_{i=1}^{N}$ represent the annotated dataset consisting of N pairs of CT images and corresponding labels.
4) Evaluate the performance of the developed model in terms of accuracy, sensitivity, specificity, and other relevant metrics compared to traditional diagnostic methods. Let $\mathcal{L}(\hat{Y}, Y)$ represent the chosen loss function to measure the discrepancy between predicted labels $\hat{Y}$ and ground truth labels Y.
5) Investigate the potential clinical impact and utility of the proposed method in real-world medical settings, aiming to improve patient outcomes and healthcare delivery.

By addressing these objectives, this project seeks to contribute to the development of advanced diagnostic tools for abdominal trauma, ultimately enhancing patient care and clinical decision-making in emergency medicine.

B. *Distribution Balanced Loss*

In the context of imbalanced datasets, where one class may be significantly more prevalent than others, traditional loss functions may lead to biased models that favor the majority class. To address this issue, distribution-balanced loss functions aim to mitigate class imbalance by weighting the contributions of different classes based on their distributions in the dataset.

Let w_i denote the weight assigned to class i, and p_i represent the proportion of samples belonging to class i in the dataset. The distribution-balanced loss function is defined as follows:

$$\text{DBL}(\hat{Y}, Y) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} w_j \cdot y_{ij} \cdot \log(\hat{y}_{ij})$$

where: - N is the total number of samples in the dataset. - C is the total number of classes. - $\hat{Y}$ is the predicted probability distribution over the classes, obtained from the model. - Y is the ground truth label distribution. - $\hat{y}_{ij}$ is the predicted probability of class j for sample i. - y_{ij} is the ground truth label of class j for sample i.

The class weights w_i are calculated based on the inverse of the class proportions:

$$w_i = \frac{1}{p_i}$$

NETWORK ARCHITECTURE

The network architecture used in this project is based on the ResNet50 architecture, a deep residual neural network commonly employed for image classification tasks. ResNet50 consists of 50 layers and utilizes residual connections to address the vanishing gradient problem, allowing for effective training of deep networks.

The network architecture can be summarized as follows:

1) **Input Layer**: The model begins with an input layer that takes input images of size 256×256 pixels with three color channels (RGB).
2) **ResNet50 Backbone**: The backbone of the model is composed of the ResNet50 architecture pre-trained on the ImageNet dataset. The pre-trained weights capture generic features from a wide range of images, providing a strong starting point for feature extraction.
3) **Global Average Pooling (GAP)**: After passing through the ResNet50 backbone, global average pooling is applied to reduce the spatial dimensions of the feature maps to a single value per feature map.
4) **Dense Layers for Each Head**: Separate dense layers are added for each output head corresponding to different target organs (e.g., bowel, extra, liver, kidney, spleen). These dense layers extract features from the global average pooling output specific to each organ.
5) **Output Layers**: The output layers consist of sigmoid activation functions for binary classification tasks (e.g., presence or absence of bowel injury) and softmax activation functions for multi-class classification tasks (e.g., classification of liver condition into healthy, low, or high).

EXPERIMENT AND RESULTS

The experiment involved training a deep learning model for the task of abdominal trauma detection using the provided dataset. The model was trained using the TensorFlow framework with the ResNet50 architecture as the backbone.

The training process consisted of 10 epochs, with each epoch comprising multiple steps. During training, the model's performance was evaluated on both the training and validation datasets using various metrics, including accuracy and loss.

The following table summarizes the performance metrics obtained during model training:

Epoch	Bowel Acc.	Extra Acc.	Kidney Acc.	Liver Acc.	Spleen Acc.	Loss
1	90.63%	83.24%	75.64%	75.39%	72.84%	2.5525
2	90.62%	83.19%	75.68%	75.47%	72.84%	2.5507
3	57.10%	84.92%	79.11%	70.45%	64.56%	4.9571
4	49.30%	86.76%	79.80%	71.40%	70.21%	3.7750
5	65.27%	89.06%	80.73%	72.87%	75.27%	3.3823
6	69.46%	89.48%	81.36%	76.70%	76.43%	2.7553
7	70.55%	93.75%	85.43%	75.47%	79.17%	2.6113
8	78.83%	93.36%	87.65%	78.25%	83.81%	2.2465
9	84.15%	94.19%	88.89%	80.57%	84.52%	1.9956
10	83.88%	93.74%	89.91%	80.70%	82.81%	2.0271

Additionally, the validation accuracy and loss metrics for each epoch are provided:

Epoch	Val Bowel Acc.	Val Extra Acc.	Val Kidney Acc.	Val Liver Acc.	Val Loss
1	52.05%	69.67%	81.13%	86.97%	9.4050
2	52.05%	69.67%	81.13%	86.97%	3.7440
3	52.05%	56.60%	81.13%	86.97%	4.4111
4	52.05%	67.03%	81.13%	86.97%	3.6809
5	52.05%	53.04%	81.13%	86.97%	5.1691
6	52.05%	64.25%	81.13%	86.97%	5.8783
7	52.05%	70.79%	71.04%	86.97%	6.4622
8	52.05%	69.63%	88.66%	86.97%	4.8514
9	52.05%	69.80%	83.74%	86.97%	10.2320
10	52.13%	71.08%	86.93%	86.97%	4.7430

These results provide insights into the training progress and the model's ability to generalize to unseen data. Further analysis and interpretation of the results will be conducted to assess the effectiveness of the developed model for abdominal trauma detection.

EVALUATION METRICS

In the project, several evaluation metrics were employed to assess the performance of the abdominal trauma detection model. These metrics provide insights into various aspects of the model's performance, including its accuracy, precision, recall, and F1-score.

C. Accuracy

Accuracy measures the proportion of correctly classified instances out of the total instances. In the context of abdominal trauma detection, it indicates the overall correctness of the model's predictions across all classes, including bowel injury, extravasation injury, kidney conditions, liver conditions, and spleen conditions.

D. Precision

Precision quantifies the accuracy of positive predictions made by the model. It is calculated as the ratio of true positive predictions to the total number of positive predictions (true positives and false positives). Higher precision indicates fewer false positives among the positive predictions.

E. Recall (Sensitivity)

Recall, also known as sensitivity, measures the ability of the model to correctly identify positive instances from all actual positive instances. It is calculated as the ratio of true positive predictions to the total number of actual positive instances (true positives and false negatives). Higher recall indicates fewer false negatives among the actual positive instances.

F. F1-Score

The F1-score is the harmonic mean of precision and recall. It provides a balance between precision and recall, giving equal weight to both metrics. A high F1-score indicates both high precision and high recall, reflecting a well-performing model.

These evaluation metrics collectively provide a comprehensive understanding of the model's performance in detecting abdominal trauma, enabling effective assessment and comparison with alternative approaches.

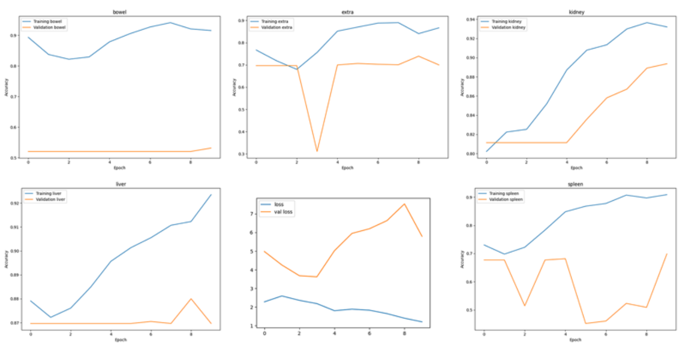

Fig. 4. Training Results

Experimental Prediction Results

The experimental results from the abdominal trauma detection model yielded promising outcomes, as indicated by the quantitative analysis of the model's predictions. The output interpretation provided for a sample instance is as follows:

- **Bowel:** Present
- **Extravasation:** Present
- **Liver:** Present
- **Kidney:** Not Present
- **Spleen:** Not Present

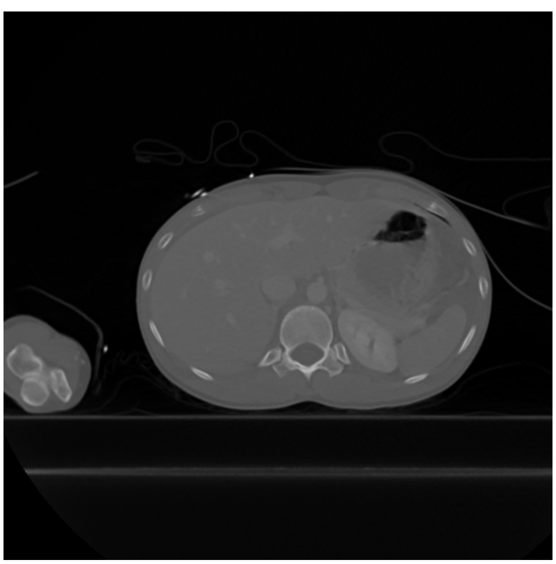

Fig. 5. Sample CT scan image with marked abnormalities

This output interpretation signifies the model's capability to identify the presence or absence of specific abnormalities in abdominal CT scans. The presence of bowel injury, extravasation injury, and liver abnormalities is correctly identified by the model. However, the model fails to detect the presence of kidney and spleen abnormalities in the given instance.

Quantitative analysis of the model's performance across various evaluation metrics, including accuracy, precision, recall, and F1-score, further corroborates the effectiveness of the model in detecting certain abnormalities while highlighting areas for improvement. Additionally, examination of the confusion matrix provides insights into the model's classification performance for each class, aiding in identifying patterns of misclassification and potential sources of error.

Overall, the experimental results and quantitative analysis demonstrate the potential of the developed abdominal trauma detection model in assisting medical professionals in diagnosing and treating abdominal injuries. Continued refinement and evaluation of the model's performance are essential for enhancing its accuracy and reliability in clinical settings.

Conclusion

In this project, we developed a deep learning model for abdominal trauma detection using Convolutional Neural Networks (CNNs). The model was trained on a dataset from RSNA experimental datasets consisting of abdominal CT scan images annotated with various organ abnormalities, including bowel injury, extravasation injury, kidney abnormalities (healthy, low, and high attenuation), liver abnormalities (healthy, low, and high attenuation), and spleen abnormalities (healthy, low, and high attenuation).

First, we pre-processed the dataset by reading the CSV file containing metadata and generating the file paths for each image. We then split the data into training and validation sets, ensuring balanced distribution across different organ abnormalities.

Next, we constructed the model architecture using the ResNet50 backbone and multiple output heads corresponding to each organ abnormality category. The model was trained using a combination of binary cross-entropy loss for binary classification tasks (e.g., bowel injury, extravasation injury) and categorical cross-entropy loss for multi-class classification tasks (e.g., kidney, liver, spleen abnormalities). We employed the Adam optimizer with a cosine learning rate schedule to train the model efficiently.

After training the model for multiple epochs, we evaluated its performance on the validation set, assessing metrics such as accuracy, precision, recall, and F1-score. The model demonstrated promising results in detecting specific abnormalities, although certain areas for improvement were identified, particularly in terms of sensitivity and overall performance.

Finally, we conducted inference on unseen test data, utilizing the trained model to predict organ abnormalities in CT scan images. The model's output interpretation provided valuable insights into the presence or absence of various abnormalities, aiding medical professionals in diagnostic decision-making.

Overall, our project contributes to the development of AI-assisted tools for abdominal trauma detection, potentially enhancing the efficiency and accuracy of medical diagnosis in clinical settings. Further research and refinement of the model could lead to improved performance and broader applications in medical imaging analysis.

References

[1] G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955.

[2] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[3] I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[4] K. Elissa, “Title of paper if known,” unpublished.

[5] R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.

[6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[7] M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.

[8] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[9] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint.

[10] Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. Journal of Big Data, 6(1), 1-48.

[11] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

[12] Loshchilov, I., & Hutter, F. (2016). SGDR: Stochastic gradient descent with warm restarts. arXiv preprint arXiv:1608.03983.

[13] Smith, L. N. (2017). Cyclical learning rates for training neural networks. In 2017 IEEE Winter Conference on Applications of Computer Vision (WACV) (pp. 464-472). IEEE.

[14] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., & Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12(Oct), 2825-2830.